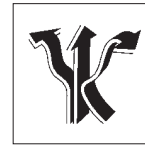
МІЖРЕГІОНАЛЬНА
АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ



МАУП

БОЛОНСЬКИЙ ПРОЦЕС: реалізація положень Болонської декларації в системі вищої освіти України

Матеріали V Міжнародної наукової конференції,
м. Судак, Автономна Республіка Крим,
20–22 вересня 2005 р.

Київ 2006

ББК 74.58(4УКР)я43
Б79

Відповідальний редактор *В. М. Чирков*
(видання здійснене в рамках проекту “Наукові праці МАУП”)

Б79 **Болонський процес: реалізація положень Болонської декларації в системі вищої освіти України:** Матеріали V Міжнар. наук. конф., м. Судак, Автономна Республіка Крим, 20–22 верес. 2005 р. — К. : МАУП, 2006. — 208 с.: іл. — Бібліогр. в кінці статей.

ISBN 966-608-597-6

У пропонованому збірнику наукових праць висвітлюються проблеми методології і методики дослідження особливостей становлення європейського освітнього простору в контексті Болонського процесу. Досліджуються і аналізуються стратегічні, тактичні та технічні аспекти входження України до європейського і світового освітнього простору, а також проблеми вдосконалення стандартів національної вищої школи і запровадження сучасних освітніх технологій.

Для науковців, державних службовців, менеджерів освіти, викладачів і студентів, які навчаються за такими напрямами і спеціальностями: державне управління, політологія, менеджмент, маркетинг, педагогіка, міжнародні відносини.

ББК 74.58(4УКР)я43+74.58я43

ISBN 966-608-597-6

© Міжрегіональна Академія управління персоналом (МАУП), 2006

Розділ ІІ
УДОСКОНАЛЕННЯ ПЕДАГОГІЧНОГО ПРОЦЕСУ У ВИЩІЙ ШКОЛІ УКРАЇНИ 84

Кузнєцова О. О.
Шляхи вдосконалення педагогічного процесу у вищій школі 85

Семаков Г. С.
Забезпечення та гарантії якості вищої освіти в контексті Болонської декларації 92

Бабкіна О. В.
Проблеми управління якістю підготовки спеціалістів у контексті Болонських реформ 97

Корнєєва Л. О., Сініцький М. Є.
Впровадження тестових методів оцінювання успішності навчання як напрям інтеграції вітчизняної освіти в Європейський освітній простір 104

Акмаєв А. І., Долголаптєв В. М., Петрушов С. М., Шевцов Л. В.
Використання нових технологій у навчальному процесі як інструменту підвищення якості технічної освіти 117

Петрушов С. М., Шевцов Л. В.
Роль самостійної роботи студента при організації навчання з використанням кредитно-модульної системи 119

Шегедин М. Б.
Шляхи реформування вищої базової медичної освіти в контексті Болонської декларації 121

Медведєв В. К., Стасєв Ю. В.
Проблеми впровадження основних положень Болонської декларації в систему військової освіти в Україні 127

Міхатуліна О. М.
"Юридична клініка" як інноваційна форма навчального процесу в підготовці юристів 133

Чобан М. В.
Некоторые проблемы преподавания английского языка как всемирного средства общения 140

Баєва О. В.
Впровадження європейського досвіду при підготовці менеджерів у галузі охорони здоров'я 143

Новальська Н. І.
Досвід впровадження сучасних освітніх технологій при викладанні навчальної дисципліни "Основи менеджменту" 151

Л. О. КОРНЄЄВА
зав. лабораторією
(Національний аграрний університет, м. Київ)
М. Є. СІНИЦЬКИЙ
канд. фіз.-мат. наук,
заст. директора Вищої школи управління МАУП

ВПРОВАДЖЕННЯ ТЕСТОВИХ МЕТОДІВ ОЦІНЮВАННЯ УСПІШНОСТІ НАВЧАННЯ ЯК НАПРЯМ ІНТЕГРАЦІЇ ВІТЧИЗНЯНОЇ ОСВІТИ В ЄВРОПЕЙСЬКИЙ ОСВІТНІЙ ПРОСТІР

Серед проблем, з якими доведеться мати справу професорсько-викладацькому складу вищих навчальних закладів (ВНЗ) України на шляху впровадження ідей Болонської декларації, є застосування у практиці поточного та підсумкового педагогічного контролю (у тому числі вступних та державних іспитів) тестових методів визначення рівня знань, умінь, навичок та інших світоглядних характеристик абітурієнтів і студентів.

Чому саме тестовим методикам уже понад півстоліття віддають перевагу навчальні заклади Європи та Америки над звичайними усними та письмовими формами перевірки знань, а в багатьох педагогічних ВНЗ (у тому числі на теренах СНД) викладають дисципліну, що має назву "тестологія"?

Тести виникли на початку XX ст. як емпіричні методи психології, що на відміну від звичайних загадок, шарад, головоломок мають за мету визначення наявності та вимірювання певних індивідуально-психологічних характеристик людини. Розрізняють тести загального інтелекту, що представляють певні розумові завдання, за результатами виконання яких визначають деяку інтегральну оцінку розвитку людини, і тести особистості, що дають змогу диференціювати особистісні прояви певної властивості [4].

Поступово психодіагностичні методики були використані для оцінювання рівня здобутих знань, умінь та навичок школярами і сту-

дентами, а згодом посіли вагоме місце в західних педагогічних системах [3]. Сьогодні у США, наприклад, тестування використовується для відбору абітурієнтів на відповідні програми навчання, з метою визначення сильних і слабких сторін студентів, для вступного, поточного та підсумкового контролю (у тому числі у ліцензійних та сертифікаційних іспитах), з метою інтерактивного навчання та корекції індивідуальних навчальних планів.

Щоб визначити переваги того чи іншого методу педагогічного вимірювання, треба порівняти їх можливості задовольняти вимогам, що звичайно висувають до якості засобів педагогічної діагностики та вимірювання:

- об'єктивності (справедливості);
- валідності (адекватності);
- надійності (достовірності);
- точності.

Переваги тестових методів перед звичайними опитуваннями та іспитами в *об'єктивності* є загальновизнаними і пояснюються декількома факторами:

- відсутністю безпосереднього спілкування викладача зі студентом або неучастю викладача в інтерпретації результатів, що виключає можливість пристрасного оцінювання та захищає студента від типових помилок екзаменаторів: поблажливості або суворості (емоційного суб'єктивізму); штучного занижування варіабельності оцінок; орієнтації на рівень підготовки академічної групи (в добре підготовлених групах екзаменатори доволі часто занижують оцінки через те, що намагаються відтворити нібито наявну схему розподілу оцінок) та ін.;
- значно більшим обсягом контролю (кількістю тестових завдань), ніж у традиційному іспиті, коли викладач змушений оцінювати рівень засвоєння студентом всієї навчальної програми за відповідями на невелику кількість питань. Імовірність помилитися в цьому варіанті значно більша, ніж при тестуванні, особливо, коли екзаменатор не працював зі студентом як викладач;
- кращими психологічними умовами та відповідно емоційним станом студента при тестуванні. Встановлено, що екзаменаційна оцінка на традиційному усному іспиті здебільше залежить не тільки від рівня спаравжніх знань студентів, а й від їхнього вміння послабити тривожний стан, вміння пристосуватися до

екзаменаційного напруження, від особистих стосунків з викладачем, уміння спілкування та інших суб'єктивних факторів. В умовах тесту дія цих факторів слабкіша, а ступінь уважності студента, його здатність до логічних висновків вищий, ніж за безпосереднього спілкування з екзаменатором;

- у тестових випробуваннях правильна відповідь на кожне питання тесту встановлюється заздалегідь колегіально групою розробників тесту, тобто є об'єктивною;
- тести та умови їх проведення можуть бути стандартизовані, що надає можливості поєднувати результати та використовувати методи математичної статистики для обробки сукупних масивів інформації, інтерпретувати результати на підставі загальновизнаних алгоритмів, прозорих правил та єдиних підходів;
- тестові методи реалізуються на ЕОМ, тобто процеси накопичення, збереження та обробки даних тестування максимально об'єктивні і ніяк не упереджені.

Під *валідністю* методів психологічного чи педагогічного контролю розуміють ступінь відповідності між тим, що має вимірюватись, і тим, що фактично вимірюється. У психодіагностиці розрізняють декілька видів валідності: зовнішню та внутрішню, конструктну, конвергентну та дискримінантну, практичну і прогностичну та ін. [4]. Валідність методів вимірювання успішності засвоєння знань та вмінь їх застосовувати поділяють на валідність змісту, відповідності та прогнозу [1].

Зовнішню валідність (очевидну) характеризують враженням, яке утворюється у неспеціаліста від ознайомлення зі змістом випробування. Вона практично не має відношення до істинної валідності як кількісної характеристики, але виконує дуже важливу допоміжну функцію — характеризує довіру чи недовіру до змісту випробування. Очевидно, якщо той, хто екзаменується, вважатиме питання недоречними, неадекватними, хибними, навряд чи розкриє цілком свої знання. Високий рівень зовнішньої валідності — це, за мовою комп'ютерних наук, якісний інтерфейс між методикою та користувачем. Зрозуміло, що багатогранність та більша змістовність тестів надає кращі можливості встановити контакт з випробуваним, досягти високого ступеня довіри до тестів, ніж до викладача, який не має необхідного авторитету.

Валідність змісту — це відповідність того, що вимірюється, змісту навчання, тобто навчальному плану та переліку необхідних знань,

вмінь та навичок, що студент повинен отримати у процесі навчання. Від валідності змісту здебільше залежить достовірність результатів педагогічного вимірювання. За цією характеристикою тестові методи також можуть мати переваги перед традиційними за рахунок кращої проробки та глибини і всебічності питань.

Валідність відповідності — це відповідність результатів вимірювання однієї ознаки різними методами. В цьому плані, на нашу думку, на етапі впровадження тестових методик мають співіснувати як тестові, так і звичайні методики педагогічного контролю. Дуже добре, якщо питання для тестів розроблятимуть незалежні фахівці, у тому числі іноземні. Статистичне порівняння отримуваних результатів надає можливості розробити як еталони тестових методик місцевого рівня, так і міжнародного ґатунку, а це є шляхом до визнання якості освіти, отриманої в різних країнах, до чого, власне, і закликає Болонська декларація.

У системі стандартів вищої освіти України в якості статистичного показника валідності змісту та відповідності тестових методик пропонують використовувати коефіцієнт кореляції між результатами вимірювання методом, що апробується, та еталонним. Але за відсутності останнього поки що цей показник на практиці не використовується.

Валідність прогнозу, або прогностична валідність — це ступінь відповідності прогнозованої успішності певної діяльності такій, що відбувається насправді. Оскільки валідність прогнозу досліджується у часі, то її рівень безпосередньо залежить від стабільності вимірюваної властивості. Якщо знання та вміння не є глибокими і сталими, валідність прогнозу визначається не точно. За результатами звичайних усних іспитів визначити валідність прогнозу досить складно, оскільки умови випробувань надто варіабельні (важко набрати статистику хоча б зі 100–200 спостережень, не змінюючи склад приймальної комісії тощо).

Надійність методу педагогічного вимірювання — це ступінь стійкості отримуваних результатів при повторних вимірюваннях властивостей тих самих осіб за допомогою того самого методу. Надійність визначає діапазон розсіювання оцінок, отриманих у декількох повторних випробуваннях. Кількісною характеристикою цієї величини є коефіцієнт надійності, що розраховують на основі виміру коефіцієнта кореляції Пірсона між результатами повторних випробувань, виконаних в ідентичних умовах за формулою Спірмена-Брауна:

$$R=k*r/[1+(k+1)], \quad (1)$$

де k — кількість частин, на які був поділений тест; r — коефіцієнт кореляції Пірсона.

Тут, як і у випадку з валідністю прогнозу, результат залежить як від характеристик засобу вимірювання, так і від стабільності у часі властивості, що вимірюється. Але тут, на відміну від процедури прогнозування, можна використовувати паралельні спостереження, тобто ставити людині питання по суті однакові за змістом, але різні за формою. Таку можливість, безумовно, надає тестовий метод, водночас дуже проблематично її використання на усних іспитах. При комп'ютерному тестуванні з використанням мережевих технологій розрахунки статистичних параметрів, у тому числі коефіцієнта надійності, можуть виконуватися в реальному часі, тобто студент може отримати об'єктивну і надійну оцінку майже одночасно із завершенням тестування.

З надійністю метода вимірювання нерозривно пов'язана його точність або помилка результату вимірювання. Використовуючи t-критерій (Стьюдента) для результатів, розподілених за нормальним законом, можна визначити необхідну кількість питань (довжину тесту), які потрібно поставити тому, хто проходить тестування, щоб з певною ймовірністю (звичайно 0,95) гарантувати, що випадкова складова помилки вимірювання не буде виходити за певні межі (на державних іспитах — не більше 5 %). При цьому систематична складова помилки вимірювання повинна бути визначена шляхом апробації методики на еталонних об'єктах. Такими об'єктами можуть бути масиви відповідей, в які контрольовано внесені певні помилки. Експериментально доведено, що для визначення ступеня успішності навчання з помилкою 5 % довжина тесту повинна складати 400 завдань [3]. Точність вимірювання за допомогою класичних форм перевірки знань (усних та письмових) у кращому разі досягає 25 %.

Безумовно, тестування має і певні недоліки. Складання великої кількості завдань, представлених різними способами таким чином, щоб вони були спрямовані на виявлення конкретних знань та умінь, необхідних для успішного виконання певних видів діяльності, — доволі складна задача, що вимагає спеціальних знань з тестології, психології, дидактики, математичної статистики. На жаль, чітких інструкцій стосовно написання тестів поки що немає. Не створено

також єдиної їх класифікації. Автори [3] визначають наступні складові специфікації тесту:

- мета (розподіл кандидатів на групи, поточний контроль, діагностика, корекція навчальних планів, підсумковий контроль);
- зміст, аспекти діяльності (елементи компетентності) в галузі оцінювання (домені);
- формат тесту (бланковий, комп'ютерний, лабораторний, OSCI);
- формат завдань;
- формат відповідей (закритий, напіввідкритий, відкритий, динамічний, за допомогою органів керування);
- критерії та процедури оцінювання;
- психометричні показники;
- інструктаж.

Ключовим моментом процедури розробки тестів є вибір адекватного завданням діяльності формату завдань (критеріально-орієнтованого [3]), що дає змогу об'єктивно вимірювати ступінь засвоєння конкретних навчальних елементів та надбання відповідних умінь, особливо з когнітивного домену (здатності критично мислити, вміння аналізувати та узагальнювати, приймати рішення у стандартних та складних ситуаціях). Для цього треба визначитись:

- з типом тесту: бінарний вибір, вибір "один з багатьох" (з класифікацією), множинний вибір (з класифікацією, з кумуляцією, незакінчене ствердження, порівняння та протиставлення, логічні пари), упорядкування об'єктів, маніпулювання об'єктами, побудова об'єктів, завдання з пропусками та доповненнями, коротка відповідь, відповідь-мікротвір, уведення вільно створюваного об'єкта з контролем за ключовими словами тощо;
- зі способом подання завдань (видом стимуляції): у вигляді питань, стверджень, завдань, що представлені словами — вербальних, або малюнків, схем, відео-, аудіозаписів — невербальних;
- з порядком пред'явлення стимулів: фіксованим, змінним, стохастичним, зі зворотним зв'язком.

У європейських стандартах основним типом тестових завдань є завдання множинного вибору. Їх поділяють на певні формати [3]. Проекти стандартів Міносвіти [1] містять рекомендації, за допомо-

гою яких можна побудувати схему вибору типу тесту відповідно до групи навчальних елементів, рівень сформованості знань якої має визначатися (табл. 1).

Таблиця 1

Найпопулярніші типи базових тестів та їх призначення

Тип тесту		Групи навчальних елементів (дидактичних одиниць), для вимірювання ступеня засвоєння яких використовується тест			
		Поняття	Явища	Відношення	Алгоритми
Закриті тести	Бінарний вибір	×			
	Вибір "один з багатьох"	×			
	Множинний вибір	×	×		
	Упорядкування об'єктів	×	×	×	
	Маніпулювання об'єктами	×	×	×	
	Побудова об'єктів	×	×	×	×
Відкриті тести	Завдання з пропусками та доповненнями	×	× ×		
	Завдання з короткою відповіддю	×	×	×	
	Введення вільно створюваного об'єкта з контролем за ключовими ознаками	×	×	×	×

Оскільки тестування має визначати не тільки рівень засвоєння певних знань, а й уміння їх адекватно використовувати, набір, форми та послідовність пред'явлення тестових завдань можуть слідувати логиці певних дій (ситуаційному сценарію), що той, хто складає іспит, має здійснювати, приймаючи рішення в певній проблемній ситуації. Якщо побудова тесту відповідає цим умовам, то такий тест належатиме до так званих "ситуаційних" тестів. Це найрозвиненіші тести, вони являють собою комбінації базових тестів, за їх допомогою діагностують ступінь засвоєння складної діяльності, її адекватність щодо вирішення типових задач.

Певним недоліком тестових методів є можливість вгадування правильної відповіді та наявність елемента підказки у вибіркових відповідях. Така вада насамперед властива завданням з бінарним ви-

бором відповіді, тому перевага надається тестам "один з багатьох" (за наявності чотирьох альтернативних відповідей імовірність угадування 45 % завдань в тесті довжиною 100 завдань складає 6,75*E-6) та ще більше — тестам множинного вибору з включенням так званих "дистракторів" (правдоподібних відповідей, які важко відрізнити від правильних). Дистрактори підвищують розподільну можливість тестів та зменшують імовірність угадування. В будь-якому разі тестові питання не повинні містити внутрішніх протиріч, щоб той, хто проходить тестування, займався не пошуком алогізмів, а використовував придбані знання для вибору правильної відповіді.

Як визначено проектом стандарту вищої освіти "Засоби діагностики рівня освітньо-професійної підготовки. Тести об'єктивного контролю рівня освітньо-професійної підготовки" [1], тестовий матеріал має проходити відбір у процесі його апробації (стандартизації) за результатами пілотних тестувань з метою визначення психометричних показників: коефіцієнтів надійності та валідності, стандартної помилки вимірювання, а також індексів завдань за складністю та розподільною здатністю. Останні два показники характеризують якість завдань і слугують критеріями можливості їх використання в тестах.

Однією з проблем використання тестових методик є перехід від кількості отриманих первинних балів до оцінок, тобто конвертація первинних результатів, що характеризують ступінь прояву певної якісної характеристики, у відповідне число за деякою нормованою шкалою. Справа в тому, що необхідно забезпечити можливість співставлення та сумування оцінок, одержаних в різних іспитах з різних дисциплін різними методами.

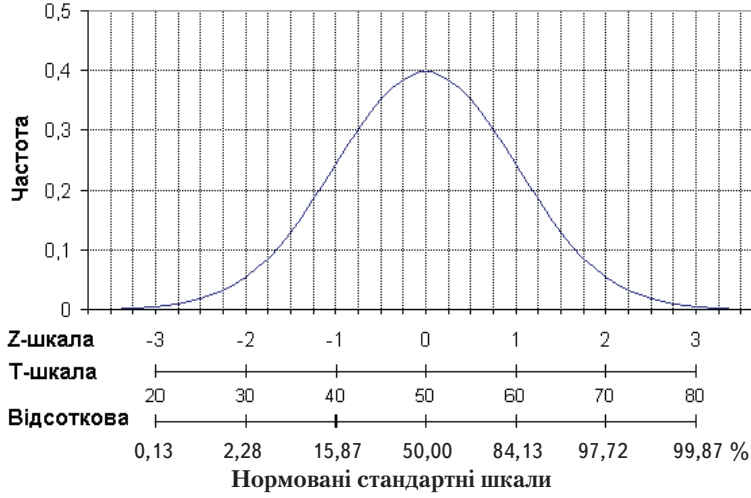
В Україні найчастіше використовують абсолютний критерій, або абсолютний стандарт підрахунку результатів, а за межу, після досягнення якої іспит вважається успішно складеним, здебільшого беруть деяку норму, наприклад, 75 % правильно виконаних завдань (оцінка "задовільно"). Це, як правило, та частина завдань, виконання яких вважається обов'язковим, і яка встановлюється a priori. Інтервал, що визначає відмінну оцінку, починається від 100 %, а його розмір визначається кількістю завдань, що дублюють інші завдання. "Добре" ділить інтервал між "задовільно" та "відмінно" на рівні частини. При цьому припускається, що всі завдання рівноцінні за складністю. Така шкала є суттєво нерівномірною і по суті ранговою, тобто такою, що не відтворює справжніх розривів між первинними результатами.

У Західних навчальних системах вже досить давно використовують відносний критерій, або відносний стандарт оцінювання. Стандарт побудований на використанні гіпотези про нормальний закон статистичного розподілу вимірюваних ознак і відповідно первинних результатів тестування за умови виконання певних умов відбору тестових завдань, що багаторазово перевірялося на вибірках національного масштабу.

Для порівняння результатів різних вимірювань, що відрізняються параметрами нормального розподілу, використовують лінійне перетворення змінної

$$Z_i = (X_i - \bar{X})/\sigma, \qquad (2)$$

тобто результати центруються відносно середнього арифметичного, а вісь абсцис градуюється у безрозмірних одиницях, які дорівнюють середньому квадратичному (стандартному) відхиленню. Таке перетворення масштабів називається нормалізацією, або стандартизацією. В результаті дані розподіляються за нормальним законом з $\bar{X} = 0$ і стандартним відхиленням $\sigma = \pm 1$ (рисунок). Згідно з правилом трьох сигм, для охоплення 99,72 % всіх результатів вистачає 6 інтервалів розміром 1σ. Побудована таким чином шкала на відміну від відсоткової є рівномірною, а її середина завжди збігається із середнім арифметичним, що представляє, як відомо, найкращу оцінку математичного очікування результату тестування.



Нормовані стандартні шкали

Конвертувати результати з використанням шкали від –3 до +3 не зовсім зручно, тому вдаються до додаткового лінійного перетворення змінних. Наприклад, у США це здійснюють за формулою

$$T_i = 50 + 10\,Z_i, \tag{3}$$

що приводить до нормального розподілу із середнім значенням 50 і стандартним відхиленням 10 (так звана *T*-шкала). Конвертована за цією шкалою величина може змінюватися в межах від 20 до 80 одиниць.

В Україні результати медичних ліцензійних іспитів конвертуються за формулою [3]:

$$T_i = 200 + 20\,Z_i, \tag{4}$$

тобто в межах ±3σ ця шкала складає від 140 до 260.

Важливо, що з даними, представленими з використанням *Z*- або *T*-шкал, можна проводити адитивні операції, у тому числі накопичувати бали для визначення рейтингу учасника за успішністю навчання або урівнювати тести, тобто конвертувати бали одного тесту в еквівалентні бали іншого. Останнє виглядає як шлях до ідентифікації якості навченості (“вагомості” оцінок) студентів, які приходять з інших закладів для продовження освіти.

Оскільки наведене справедливе тільки для нормального закону розподілу результатів тестування, існує необхідність контролю правомірності прийняття відповідної гіпотези та прийняття мір із забезпечення валідності процесу вимірювання у тих випадках, коли гіпотеза нормального закону не підтверджується. Як відомо, відхилення розподілу даних від нормального закону характеризується асиметрією As та ексцесом Ex. Аналіз знаків цих характеристик дає змогу зробити висновки відносно причин необ’єктивності вимірювання та визначити напрям валідизації (табл. 2) [3].

Таблиця 2

Критерії аналізу об’єктивності тестування

Величина асиметрії As	Величина ексцесу Ex	Висновок щодо якості тесту	Напрям валідизації
1	2	3	4
0	0	Процес вимірювання валідний	

Закінчення табл. 2

1	2	3	4
< 0	> 0	Результати завищені, розподільна здатність недостатня	Підвищити складність завдань, виключити можливість підказки
< 0	< 0	Тест легкий з великою розподільною здатністю	Підвищити складність завдань
< 0	0	Тест легкий	Підвищити складність завдань
> 0	< 0	Тест складний з великою розподільною здатністю	Зменшити кількість складних завдань
> 0	> 0	Тест складний з малою розподільною здатністю	Спростити завдання, виключити можливість підказки
> 0	0	Тест складний	Спростити завдання

Спираючись на міжнародну практику проведення ліцензійних тестувань, у проекті стандартів Міносвіти [1] пропонувалось після підрахунків суми отриманих балів (1 — за правильну відповідь, 0 — за помилкову) проводити якісний аналіз кожного завдання шляхом визначення індексів їх складності та розподільної здатності. Для цього з групи (не менше 20 осіб) вибирають третину найсильніших і третину найслабкіших результатів і визначають для кожного завдання індекс складності:

$$I_c = \frac{H + L}{n} \cdot 100\ \%, \tag{5}$$

де H та L — кількість балів, отриманих представниками відповідно сильної та слабкої груп; n — загальна кількість тих, що проходили тестування в обох групах.

Завдання, що отримали індекс складності менше 30 % або більше 70 %, є кандидатами на вилучення з розрахунків: перші як занадто складні, другі — як слабкі. Це робить тест однорідним.

Одночасно обчислюють індекс дискримінації за формулою

$$I_\partial = 2(H - L)/n. \tag{6}$$

Що більший цей індекс, то більшою мірою правильна відповідь на завдання. Якість складання завдання оцінюють за шкалою:

≥0,35 — “відмінно”; 0,25–0,34 — “добре”; 0,25–0,15 — суперечливе завдання, вимагає перегляду; < 0,15 — погане завдання, підлягає вилученню при остаточному підрахунку балів. У загальному підсумку студентам враховують бали за завдання, що мають задовільний $I_д$, а завдання з низьким I_c не включають до наступних іспитів.

Наступним питанням вимірювання якості освіти за відносним стандартом є встановлення прохідного балу. Тут використовують два критерії: відносний та абсолютний. В першому варіанті межа “склав/не склав” залежить від результатів усіх, хто здавав іспит. Наприклад, виходячи з *T*-шкали (5), можна прийняти рішення, що прохідний бал, тобто межа між оцінками “достатньо” (E) та “незадовільно” (F), становить 160 одиниць, тобто близько 2,3 % тестованих мають не здати іспит. Недолік такого підходу полягає в його певній упередженості та суб’єктивності. Але він стимулює студентів до підвищення якості знань.

Для вступних іспитів, що повинні відібрати заздалегідь визначену кількість абітурієнтів, які будуть зараховані до ВНЗ, використовують рейтингову оцінку. Вона будується як сума рангових місць (рейтингів), що абітурієнти посіли за результатами тестування з різних дисциплін. Рейтинг відповідає номеру перцентіля (1/100 *T*-шкали), в який конвертується первинний бал, і розраховується за формулою

$$R_{gi} = [100 - \text{int}(T_i - T_{\max})], \qquad (7)$$

де T_i та $T_{\max}$ — відповідно кількість балів, набраних *i*-м екзаменованим за *T*-шкалою та її максимальне значення.

На відміну від відносного критерію, прохідний бал, що розраховується за абсолютним критерієм, не залежить від розподілу результатів іспиту. Існують декілька методик визначення абсолютного критерію. Всі вони побудовані на експертних висновках щодо тестових завдань або окремих екзаменованих чи груп екзаменованих [3]. Так, за методом Нєдєльського оцінюють імовірність вгадування вірної відповіді “граничним студентом” на кожне питання і підраховують її середнє значення для тесту загалом, що і визначає прохідний бал. За методом Хофсті спочатку прогнозують максимальний і мінімальний відсотки екзаменованих, що не складуть іспит, та відповідні значення прохідного балу. Потім наносять дані прогнозу на графік залежності відсотка екзаменованих, що не склали б іспит, від величини прохідного балу, побудований за кривою розподілу попередніх (тогорічних) іспитів, і штучним прийомом визначають на цій

кривій точку, координати якої наближені до середнього прогнозованого значення обох параметрів. Ця методика вже п'ять років успішно використовується Центром тестування професійної компетентності при МОЗ України (ЦТПК МУ) для визначення величини абсолютного критерію "склав/не склав" медичних ліцензійних іспитів (МЛІ) типу "Крок". Наприклад, у 2005/2006 навчальному році для МЛІ "Крок 2. Загальна лікарська підготовка" цей критерій дорівнює 45,5 %, що відповідає 162,6 одиниць за *T*-шкалою МЛІ (5).

Тестування впродовж останніх кількох років є основним методом відбору на вступних іспитах у Києво-Могилянську академію, Львівський університет імені Івана Франка та інших вищих закладах освіти. Міністерство освіти і науки України планує у 2006 р. започаткувати тестування на випускних екзаменах у всіх середніх навчальних закладах країни.

Таким чином, накопичений досвід дає підставу для впровадження цього прогресивного напряму підвищення об'єктивності та прозорості оцінювання якості освіти у практику навчальних закладів України всіх рівнів. Саме на це, власне, і орієнтує нас Болонська декларація.

Література

1. Додаток 1 до наказу № 285 від 31 липня 1998 р. МОН України "Комплекс нормативних документів для розробки складових системи стандартів вищої освіти". — К., 1998. — 118 с.
2. Інформація веб-сайту ЦТПК МУ (www. test. ukrpack. net).
3. *Система* управління якістю медичної освіти / І. Є. Булах, О. П. Волосовець, Ю. В. Вороненко та ін. — Дніпропетровськ: АРТ-ПРЕСС, 2003. — 212 с.
4. *Скребец В. А.* Психологическая диагностика: Учеб. пособие. — 2-е изд., перераб. и доп. — К.: МАУП, 2001. — 152 с.